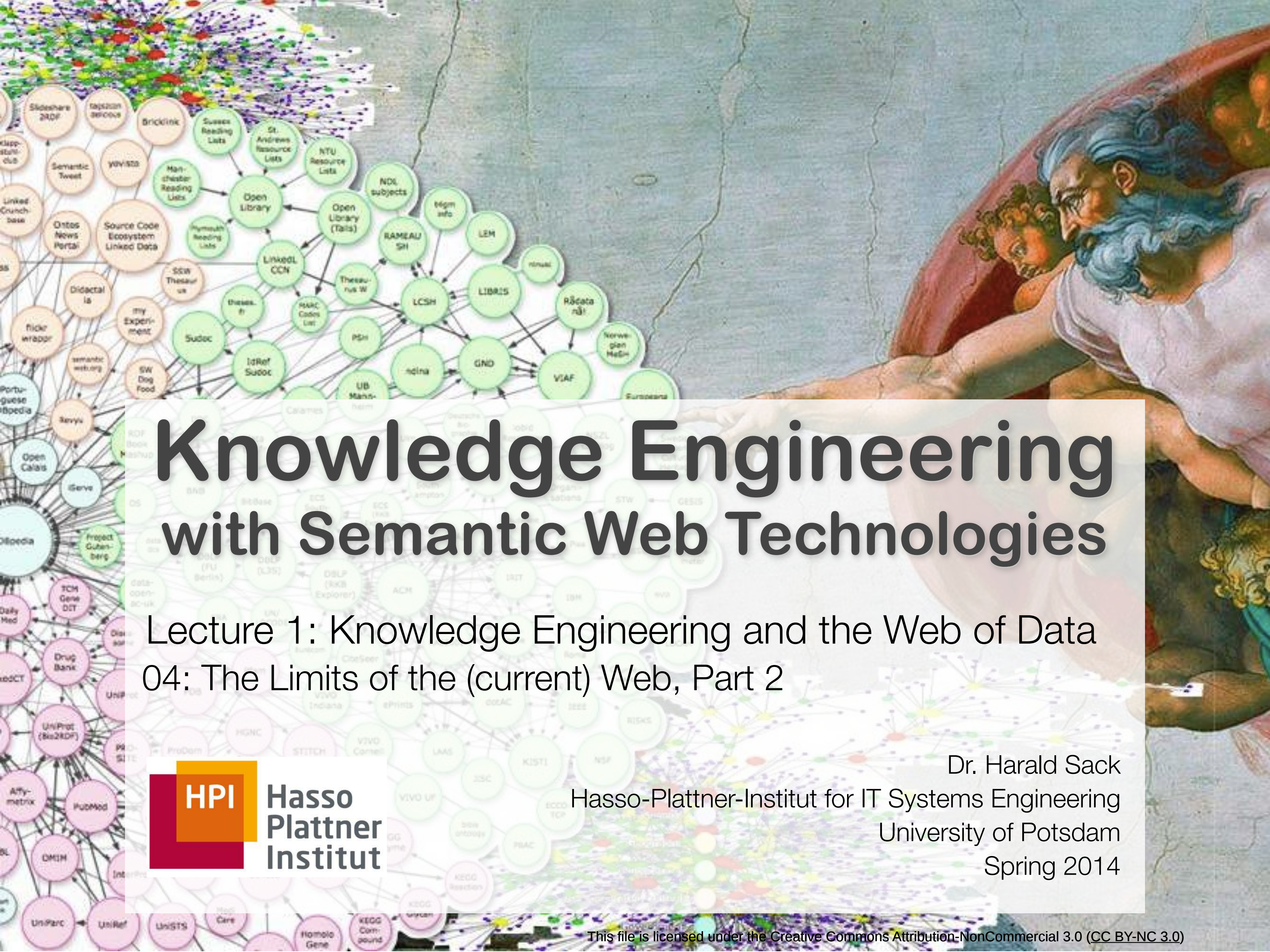




Knowledge Engineering with Semantic Web Technologies

Lecture 1: Knowledge Engineering and the Web of Data
04: The Limits of the (current) Web, Part 2



Dr. Harald Sack
Hasso-Plattner-Institut for IT Systems Engineering
University of Potsdam
Spring 2014



Lecture 1: Knowledge Engineering and the Web of Data

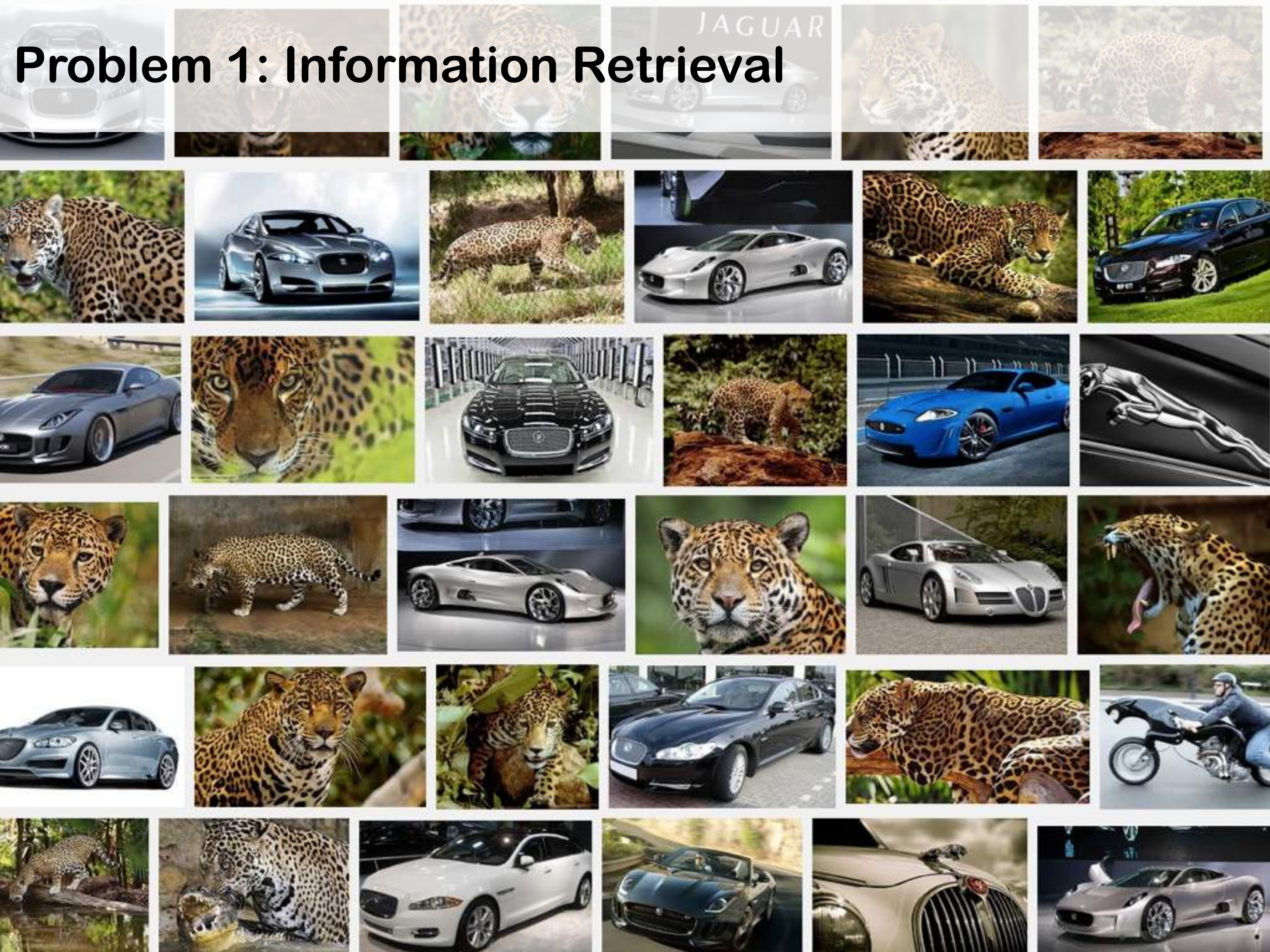
Open HPI - Course: Knowledge Engineering with Semantic Web Technologies



04: The Limits of the (current) Web Pt.2

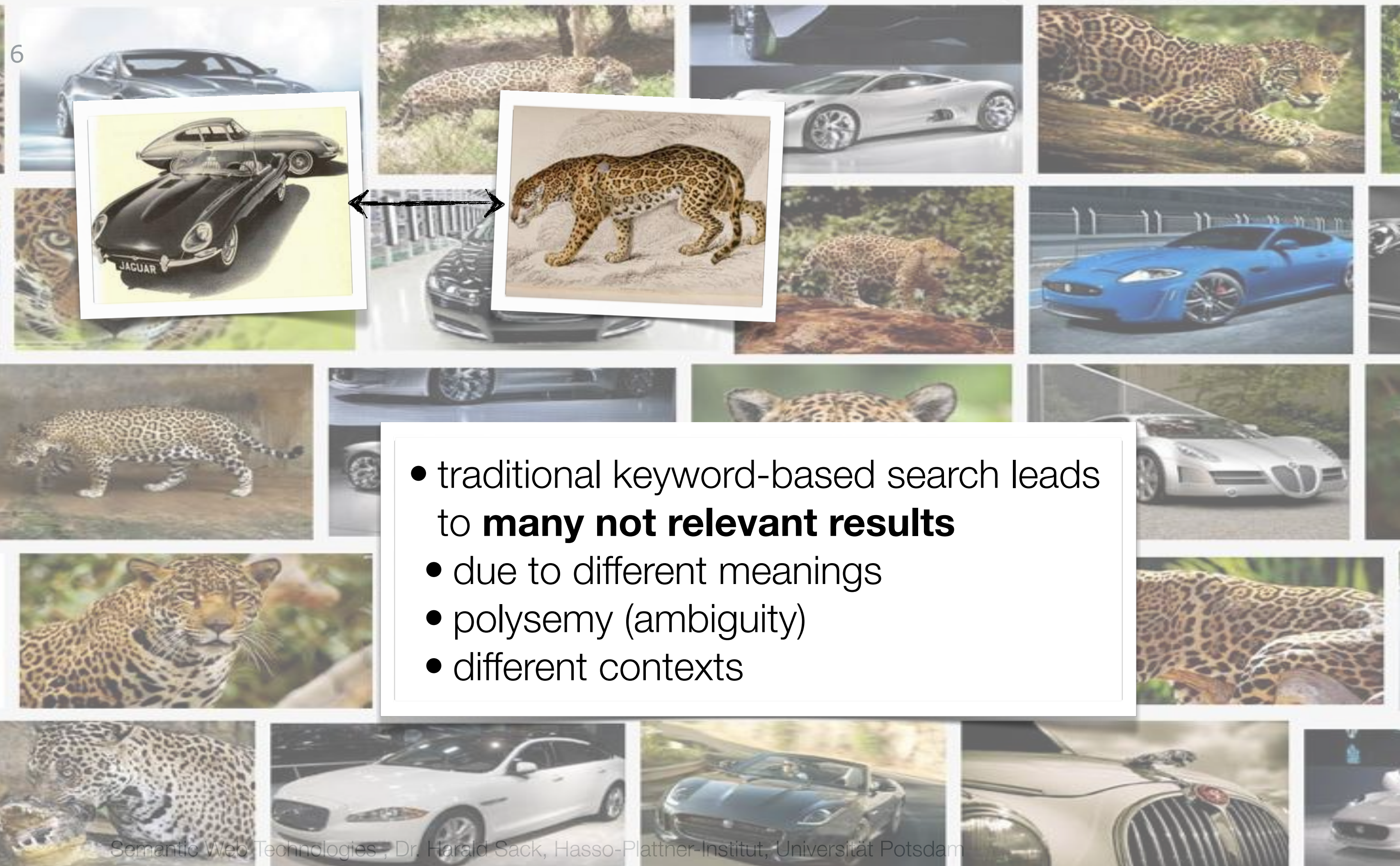
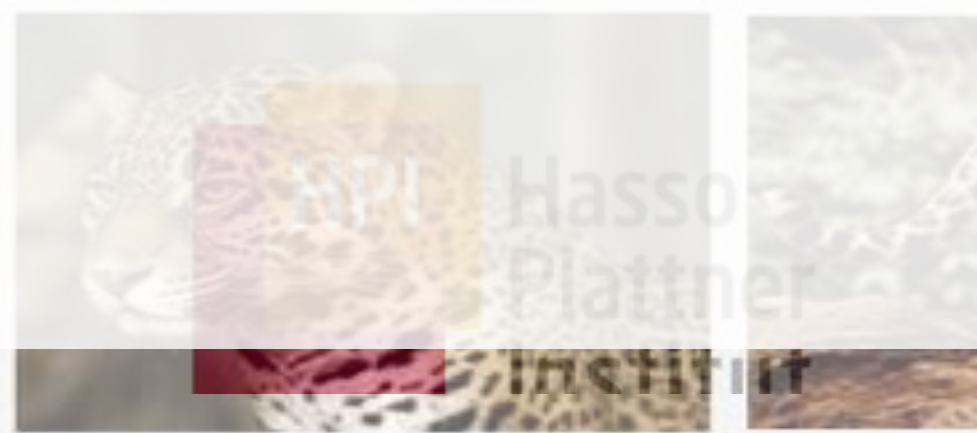
OpenHPI - Course Knowledge Engineering with Semantic Web Technologies
Lecture 1: Knowledge Engineering and the Web of Data

WHAT IS THE MEANING G



Problem 1: Information Retrieval

Problem 1: Information Retrieval



- traditional keyword-based search leads to **many not relevant results**
 - due to different meanings
 - polysemy (ambiguity)
 - different contexts

Problem 1: Information Retrieval

Ad related to "Panthera onca" ⓘ

[Lion, Tiger, Jaguar Pics - Panthera Works to Save Wild Cats](#)

www.panthera.org/

See Pictures of all the Big Cats

[Jaguar - Wikipedia, the free encyclopedia](#)

en.wikipedia.org/wiki/Jaguar

...ca, is the only extant New World member of the Panthera genus. It shows the lion, tiger, leopard, jaguar, snow leopard, and ...

...le Jaguars - Jaguar (disambiguation) - Jaguarundi

Jaguar

Panthera Onca

[Panthera onca \(jaguar\)](#)

www.iucnredlist.org/details/15953/0

by A Caso - 2011 - Cited by 10 - Related articles

Panthera onca. Status_ne_off Status_dd_off Status_vu_off Status_en_off Status_cr_off Status_ext_off

[ADW: Panthera onca: INFORMATION](#)

animaldiversity.ummz.umich.edu/accounts/Panthera_onca/

Panthera oncajaguar. By Jonathan Nogueira. Geographic Range; Habitat; Physical Description; Reproduction; Lifespan / Longevity. Behavior; Communication ...

[Panthera | Jaguars](#)

www.panthera.org/species/jaguar

The jaguar (**Panthera onca**) is the third largest big cat in the world, but it is the largest feline in the Americas.

[Jaguar \(Panthera onca\) - thebigzoo.com](#)

www.thebigzoo.com/Animals/Jaguar

The Jaguar is the largest cat in the Americas. It is the third largest big cat in the world (after the Lion and the Tiger).

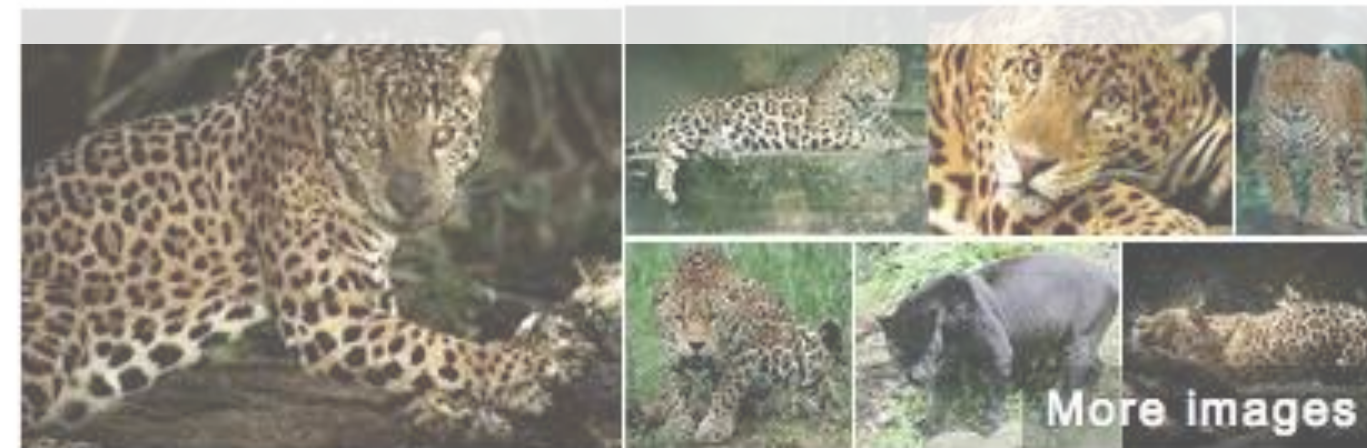
[Jaguar - Panthera onca - National Geographic](#)

www.nhptv.org/natureworks/jaguar

The jaguar is the largest cat in North America and the third largest cat in the world. It stands about three feet tall from shoulders to feet, 6½ to 7½ feet long from ...

[Jaguar videos, photos and facts - Panthera onca - ARKive](#)

www.arkive.org/Species/Mammals



Jaguar

The jaguar is a big cat, a feline in the Panthera genus, and is the only species found in the Americas. The jaguar is the third-largest big cat, after the tiger and the lion, and the largest in the Western Hemisphere.

Wikipedia

Scientific name: Panthera onca

Rank: Species

Higher classification: Panthera

[Feedback / More info](#)

- traditional keyword-based search **does not find all results**
- synonyms and metaphors
- missing context definition

Problem 2: Information Extraction

What does the information mean?

Web 東奥

県外在住者の方へ!!
東奥日報電子版
長期割引あり

新聞に購読申し込
ウェブ | 新聞

ドコモ・スマホサイト
new
ケータイ・スマホ情報

お店検索
イベント
展覧会
求人

青年海外協力隊
シニア海外ボランティア

秋募集
受付期間
10月1日～11月6日まで

JICA

青森県の天気予報
各地の気象(22時)
気温(°C) 風向/風速
青森 17.2 南南西/3m
弘前 14.0 南/2m

全国地方新聞社 秋の新商品スィーツ
50人が食べてみた! 「推し商品」はコレ!!

天地人 | 社説 | 訃報 | お悔やみ | 青森
プロ野球 | MLB | Jリーグ | 大相撲 | 高校野球
エンタメ | コラム | 新刊本 | 新作映画 | 音楽CD | 格闘技

Goウェブ

東奥日報ご愛読者限定サービス!
Web東奥の記事全文閲覧と東奥日報データベース検索 >> ログイン
東奥ウェブ読者くらぶ入会(無料)のご案内と手続き >> クリック

住まいナビ
青森県の
物件情報

東北の電力使用 21時台
71%

青森のニュース
2012年10月14日 22:37
RSS ブログパーツ

第10回 弘前・白神アップルマラソン写真特集

「滝めぐり」DVD発売
福島原発事故
東日本大震災

22:01 弘高、決勝へ／東北高校軟式野球
21:30 「生活」小沢代表が講演／十和田

不動産情報
中古車ネットガーデン
ネット住宅展示場
iメール申し込み
Miniガイド

ロレックス
正規品取扱店

マエバラ

人材派遣なら
東奥日報人材センター アウル

秋のセーティネット

サンワアリーナ 青森

風家

あかちゃん・初宮参り

クレバーホーム

Cubic Four
建築工務 キュービック・フォー

弘前初登場

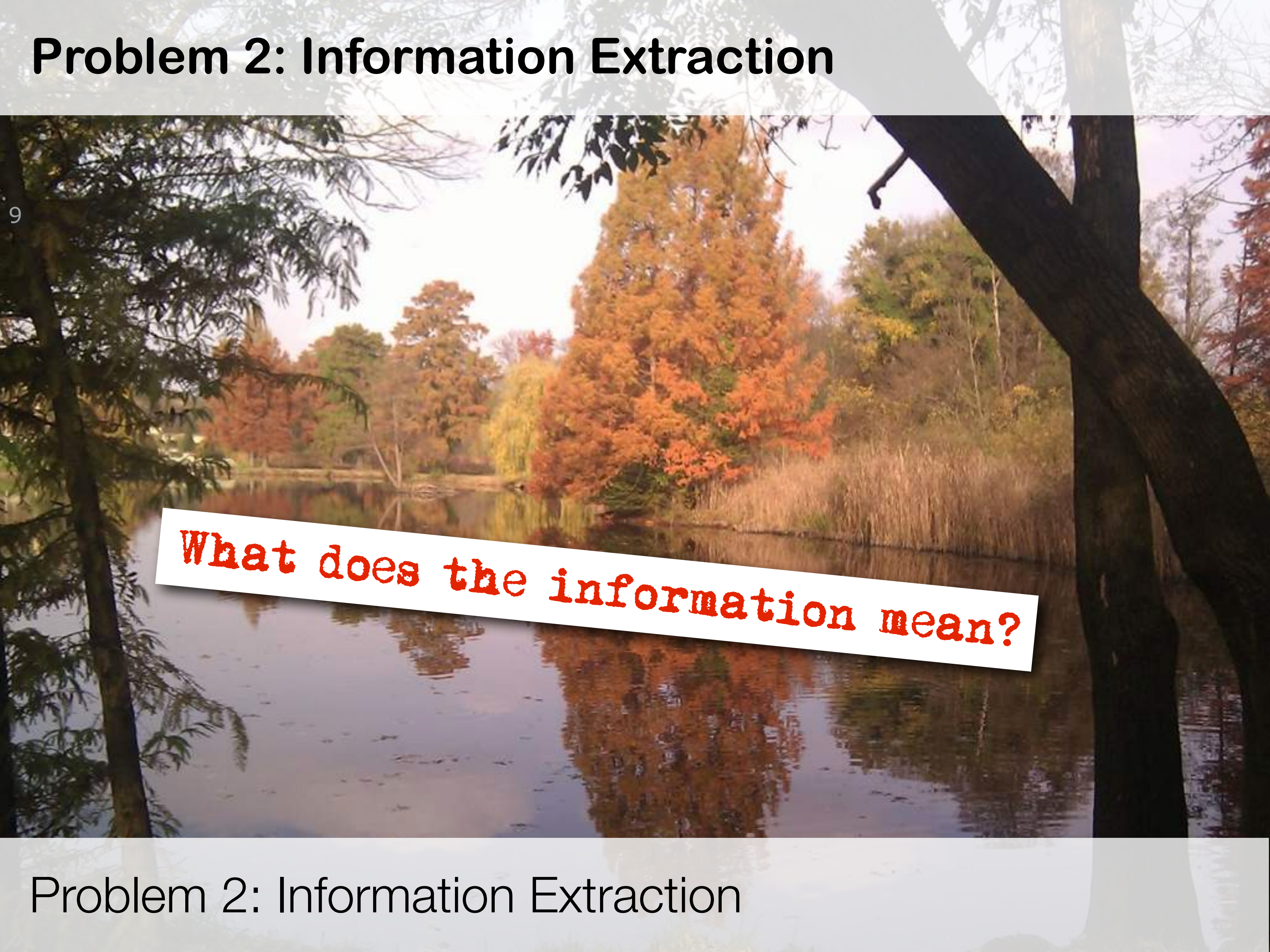
お商な情報はここから

NEW モデルハウス
オープン!!

日本瓦葺

Problem 2: Information Extraction

9



What does the information mean?

Problem 2: Information Extraction

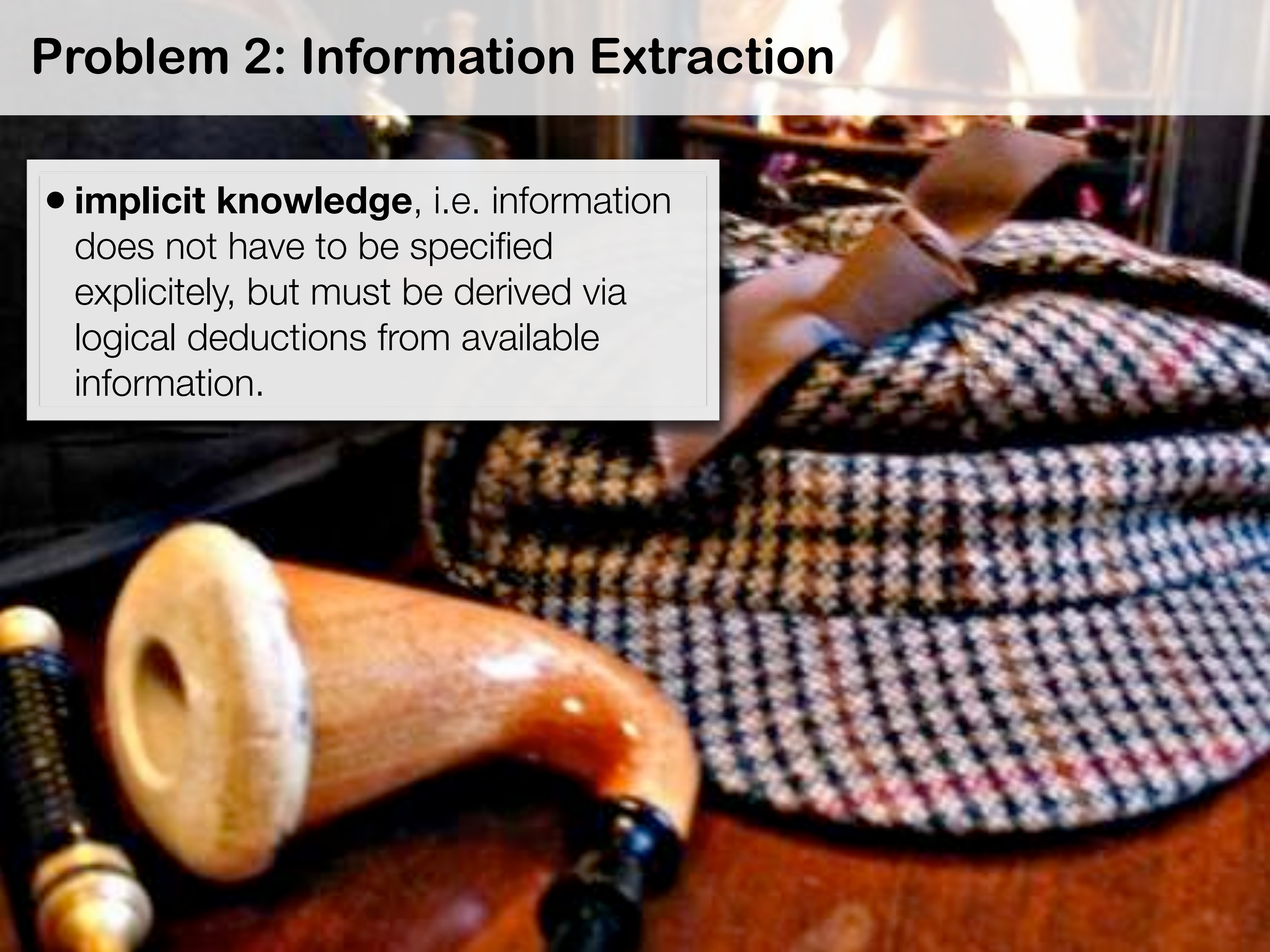
Problem 2: Information Extraction

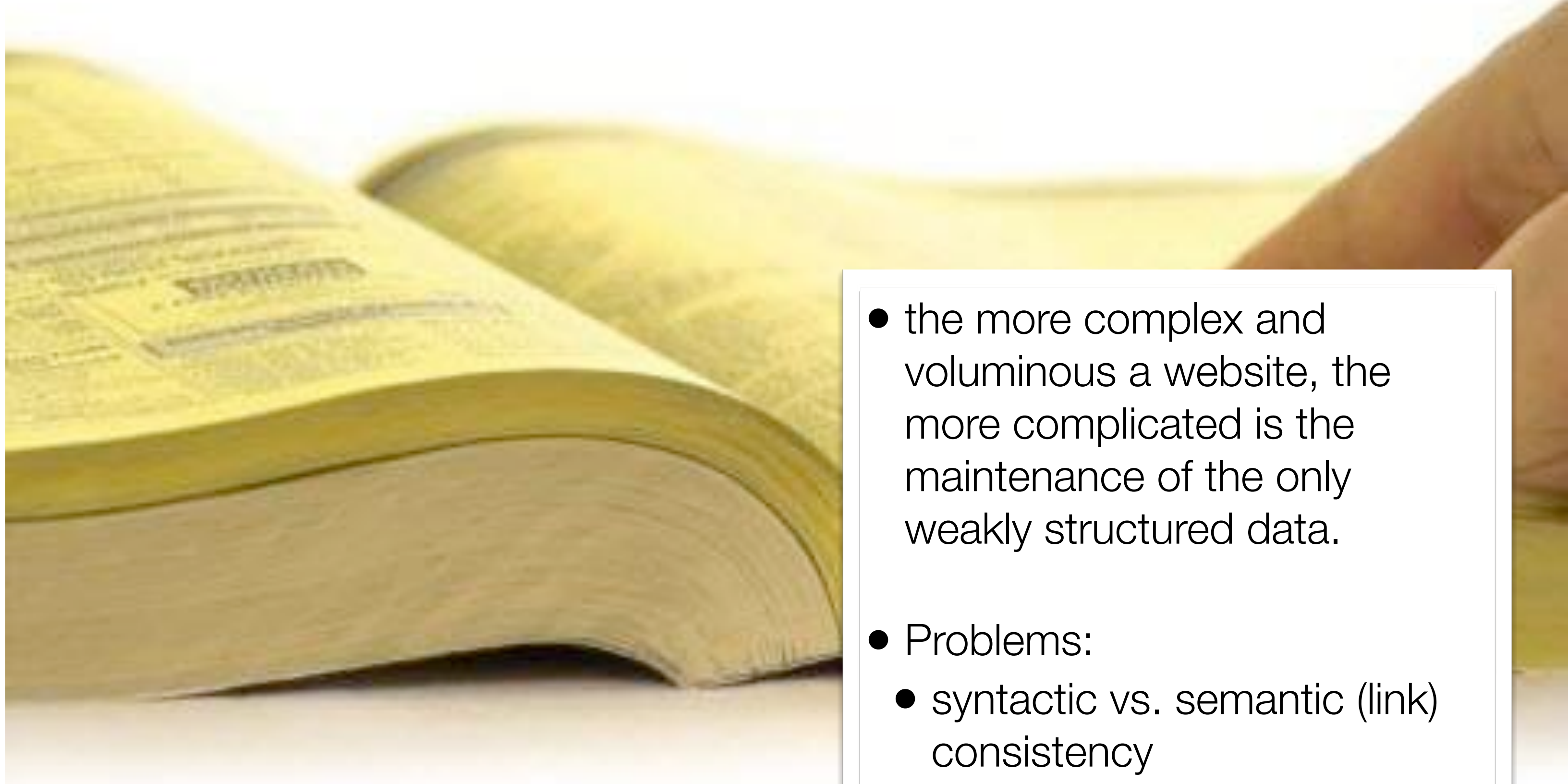
Information Extraction

- can only be solved ,*correctly*‘ by a human agent
- heterogeneous distribution and order of information
- a software agent does not have
 - sufficient knowledge of contexts
 - sufficient world knowledge and
 - sufficient experience to solve the problem

Problem 2: Information Extraction

- **implicit knowledge**, i.e. information does not have to be specified explicitly, but must be derived via logical deductions from available information.





- the more complex and voluminous a website, the more complicated is the maintenance of the only weakly structured data.
- Problems:
 - syntactic vs. semantic (link) consistency
 - correctness
 - timeliness

Problem 3: Maintenance

- Adaption of the presented information content to personal requirements
- Problems:
 - from where do we get the required (personal) information?
 - personalization vs. data security



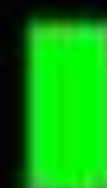
Problem 4: Personalization

SCORE 7040

LIVES



GAME OVER



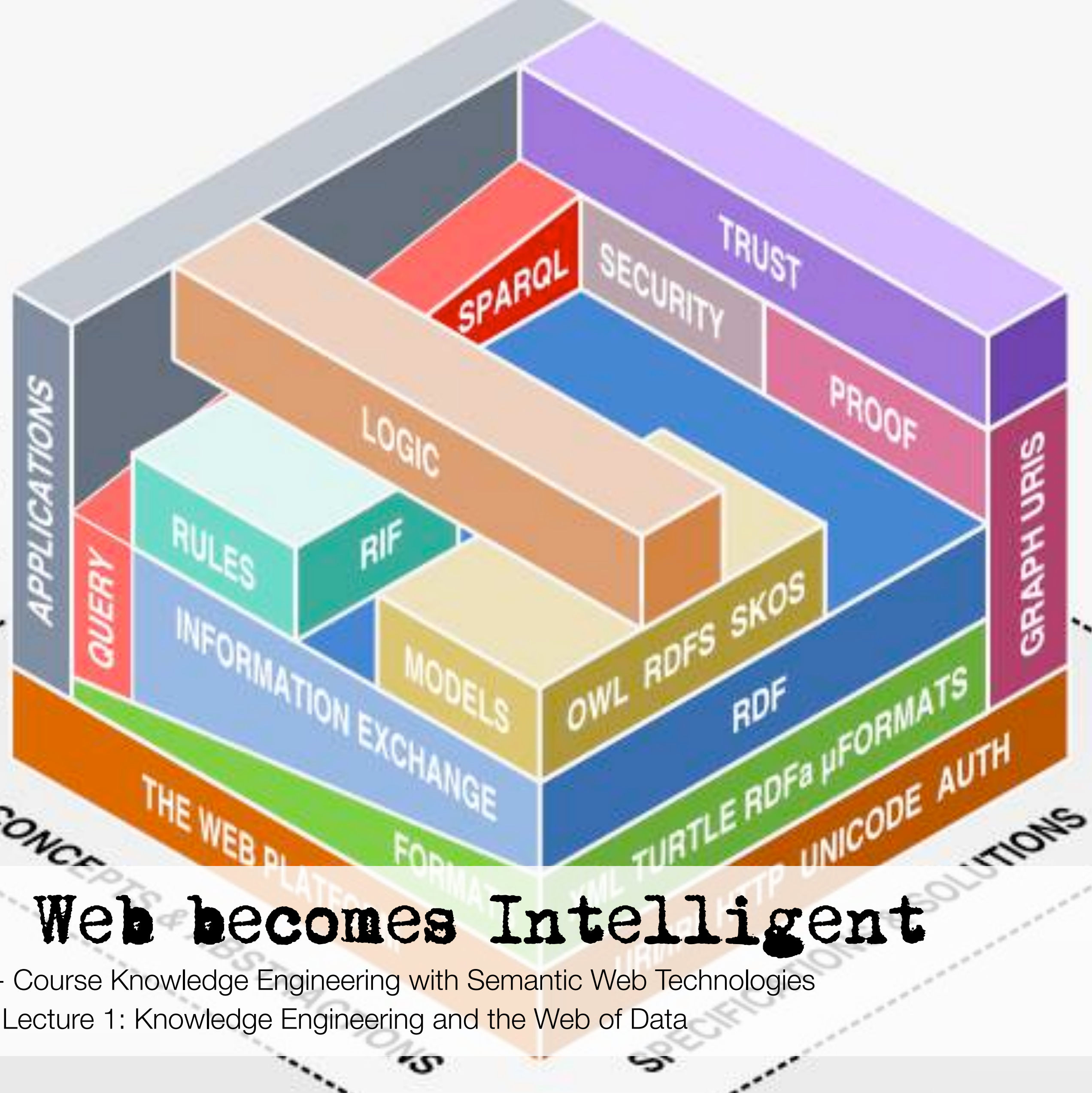
e of cake...)

Next section

15

subset of the stack
grained data access
ion exchange is key
y, but not too important
based on the Web

hall
ies



05: The Web becomes Intelligent

OpenHPI - Course Knowledge Engineering with Semantic Web Technologies
Lecture 1: Knowledge Engineering and the Web of Data